AD-A244 703

CONNECTIONIST APPROACH TO TRANSFORMATION RECOVERY USING VISUAL GRADIENT DESCENT

BY GEORGE W. ROGERS   JEFFREY L. SOLKA
DONALD R. VERMILLION   CAREY E. PRIEBE
STRATEGIC SYSTEMS DEPARTMENT

NOVEMBER 1991

DTIC
ELECTE
JAN 2 4 1992
S
D
D

92-01680

NAVAL SURFACE WARFARE CENTER
Dahlgren, Virginia 22448-5000 • Silver Spring, Maryland 20903-5000

92 1 21 008

CONNECTIONIST APPROACH TO TRANSFORMATION RECOVERY USING VISUAL GRADIENT DESCENT

BY GEORGE W. ROGERS   JEFFREY L. SOLKA
DONALD R. VERMILLION   CAREY E. PRIEBE
STRATEGIC SYSTEMS DEPARTMENT

NOVEMBER 1991

FOREWORD

This report describes an artificial neural network (ANN) designed to recover an arbitrary transformation that relates two images. This ANN computes local gradients between the boundaries of the image and its transformed copy. Using the Gaussian average of these gradients, the network acts to reverse the action of the transformation. The transformation relating the two images is thus recovered in this manner.

This work has been supported by the Office of Naval Research through the Independent Research program.

This report has been reviewed by Dr. Richard A. Lorey, Head of the Space and Ocean Geodesy Branch and James L. Sloop, Head of the Space and Surface Systems Division.

Approved by:

R. L. SCHMIDT, Head
Strategic Systems Department

Accesion For		
NTIS CRA&I		☑
DTIC TAB		☐
Unannounced		☐
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

ABSTRACT

Given an object and a copy of itself produced by an unknown two-dimensional affine transformation, a new neural network architecture has been developed that recovers this transformation by minimizing the symmetric difference between the object and the copy. This architecture performs a gradient descent in symmetric difference error space and is designated as visual gradient descent (VGD). The VGD network has applications to both two- and three-dimensional model based automatic target recognition (ATR) and image compression using iterated function systems.

CONTENTS

ILLUSTRATIONS

ILLUSTRATIONS (CONTINUED)

TABLE

NOMENCLATURE

p_{ij}	=	The intensity value at pixel (i, j). For monochrome images $p_{ij} = 0$ or 1
E_{sd}	=	The symmetric difference error measure
Λ	=	A transformation on R^2 consisting of a rotation and a 2 degree of freedom (DoF) translation
F	=	A transformation on R^3 consisting of a 3DoF rotation and a 2DoF translation
Π	=	A projection mapping from R^3 into R^2
A	=	An affine transformation on R^2
$L_o(i, j; n)$	=	The integrated image intensity at pixel (i, j) for orientation number n
$\gamma_{ijn}(t)$	=	The local gradient at pixel (i, j) for orientation number n and time t
$\Gamma_n(i\lambda, j\lambda; t)$	=	The Gaussian gradient at pixel ($i\lambda$, $j\lambda$) for orientation n and time t
$\Delta r_\lambda(t)$	=	The correction vector for point number λ, λ=1..3, at time t
$T(t)$	=	The affine transformation that takes point $p\lambda$ to point $p\lambda'$, λ=1..3, at time t
$\mu(L_I)$	=	The Lebesgue measure of the left side of the one-dimensional simple cell residing in the set I
$\mu(R_I)$	=	The Lebesgue measure of the right side of the one-dimensional simple cell residing in the set I

INTRODUCTION

Many problems of interest involve recovering a transformation that relates an object and a distorted copy of itself. This task occurs in model-based automatic target recognition (ATR) and in image compression using iterated function systems. In the ATR problem, to properly identify a target image, a set of reference models needs to be maximally aligned with a target image that has been subjected to a unknown transformation. In the case of iterated function systems (IFS), an image is covered with copies of itself obtained using affine transformations. The contractive affine transformations that achieve maximal overlap between the image and the union of the copies must be recovered in order to define the IFS.

A new neural network architecture, termed visual gradient descent (VGD), has been developed that recovers an unknown transformation by computing local symmetric difference gradients between a reference object and its transformed copy. The VGD network solves for the transformation that minimizes the global symmetric difference between the object and its copy. This process can be viewed as a viable collective computation alternative to the standard global gradient descent technique.

This report begins with a description of the general object comparison problem, with an emphasis on how various types of application can be addressed by specifying the form of the transformation relating the two objects. Next, some general background on neural network approaches to vision is provided, followed by a detailed description of the VGD network when the two objects are related by an affine transformation. Following this, a theoretical analysis is performed for the case of a one-dimensional affine transformation. The report concludes with some computer simulations that recover the affine transformation relating two squares and a discussion of some ongoing and future work that applies the VGD network.

GENERAL OBJECT COMPARISON PROBLEM

ERROR MEASURES

For the purposes of this report, monochrome M by N pixel space P is defined as $\{(i,j,k) \mid i \in \{1,2, \dots ,M\} , j \in \{1,2, \dots , N\}, k \in \{0, 1\}\}$. An image I will be some subset of pixels with nonzero k values. Since any pixel $p=(i,j,1)$ that is part of an image has a nonzero k value, we will simplify our notation and write $p=(i,j)$, or indicate this pixel's k value by p_{ij}.

Given two pixel images, a measure of the degree of similarity of the images is often needed. One such measure that has the desirable property of computational efficiency is the symmetric

difference measure. The symmetric difference measure is most simply defined for two objects A and B as

$$E_{sd} = \text{number of pixels in}[A \cup B - (A \cap B)]$$

where A is the set of pixels that comprise object A, and B is the set of pixels that comprise object B. This is a set operation that can be easily implemented for binary valued pixel images as follows.

1.   Let pixel plane **a** contain object A, and pixel plane **b** object B.

2.   Summing over all i and j, compute the sum

$$E_{sd} = \Sigma\Sigma(a_{ij} - b_{ij})^2$$

$$= \Sigma\Sigma(a_{ij}^2 - 2\ a_{ij}b_{ij} + b_{ij}^{\ 2})$$

Let $x \in A \cap B$ (both set A and set B are on) then $a_{ij}=1$ and $b_{ij}=1$ hence E_{sdx}, the contribution of x to E_{sd}, is 0

if $x \in A\backslash B$ (set A is on, set B is off) then $a_{ij}=1$ and $b_{ij}=0$ hence $E_{sdx} = 1$

likewise if $x \in B\backslash A$ (set A is off, set B is on) then $a_{ij}=0$ and $b_{ij}=1$ and $E_{sdx} = 1$

the remaining possibility then is $x \in (A \cup B)'$ where $a_{ij}=0$ and $b_{ij}=0$ giving $E_{sdx} = 0$

so $E_{sd} = \Sigma\Sigma(a_{ij} - b_{ij})^2 = \Sigma(E_{sdx}) = \text{number of pixels in}[A \cup B - (A \cap B)]$ as desired

TRANSFORMATIONS RELATING OBJECTS

The application of interest determines the transformation that relates the two pixel images that need to be compared.   In the case of two-dimensional ATR where translation, within-plane rotation, and uniform scaling is allowed, then the transformed object or target image O' is obtained from one of the reference objects O_i using a transform Λ of the form

$$\Lambda \begin{vmatrix} x \\ y \end{vmatrix} = \begin{vmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} + \begin{vmatrix} e \\ f \end{vmatrix}$$

where r is the scale factor, θ is the rotation angle, and e and f are translations. Target identification is accomplished by finding which of our reference models subjected to the above transformation can be best aligned with the target image (see Figure 1).
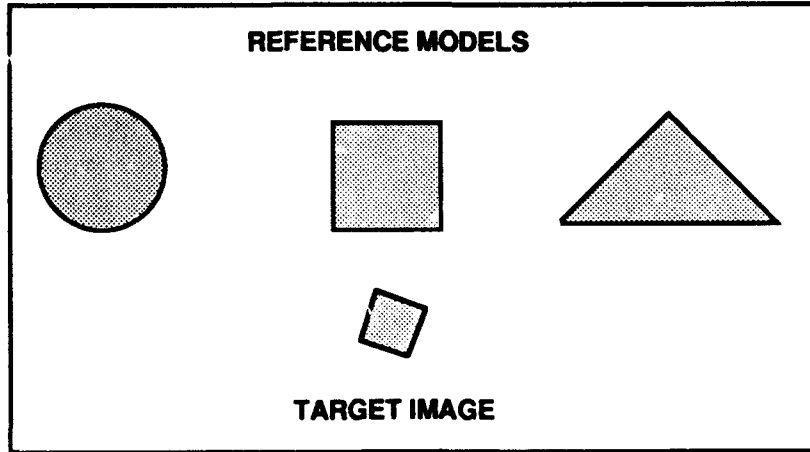
FIGURE 1. TWO-DIMENSIONAL ATR COMPARISON PROBLEM

Another problem of interest is the three-dimensional ATR problem where the target image and model images have been projected into R^2. In this case, the target image is obtained from one of the reference images via the composition of a 6 degree of freedom (DoF) mapping F with a projection mapping $\Pi : R^3 \rightarrow R^2$ into the line of sight plane. F is defined in terms of a scale factor r, x translation Δx, y translation Δy, and the angles yaw or ϕ a rotation about the z axis, pitch or θ a rotation about an intermediary y axis, and roll or ψ a rotation about the final x-axis. In this case, F is given by

$$
F\begin{bmatrix} x \\ y \\ z \end{bmatrix} = r \begin{bmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ \sin\psi\sin\theta\cos\phi - \cos\psi\sin\phi & \sin\psi\sin\theta\sin\phi + \cos\psi\cos\phi & \cos\theta\sin\psi \\ \cos\psi\sin\theta\cos\phi + \sin\psi\sin\phi & \cos\psi\sin\theta\sin\phi - \sin\psi\cos\phi & \cos\theta\cos\psi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ 0 \end{bmatrix}
$$

(Reference 1). Figure 2 portrays the three-dimensional ATR problem.

In the case of image compression using iterated function systems the two images are related by an affine transformation (Reference 2) of the form

$$
A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos\theta & -s\sin\psi \\ r\sin\theta & s\cos\psi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}
$$

where r and s are scale factors, and θ and ψ are generalized rotation angles. As part of the image compression process, an image is covered with copies of itself that are produced by affine transformations (see Figure 3). The overlap between the image and each of the collage pieces needs to be maximized for accurate compression. An analysis of a human's ability to optimally adjust the collage pieces led to the development of the VGD network. The VGD net reconstitutes the affine transformation, which optimally aligns an image with a copy of itself that has been initially subjected to an unknown affine transformation.
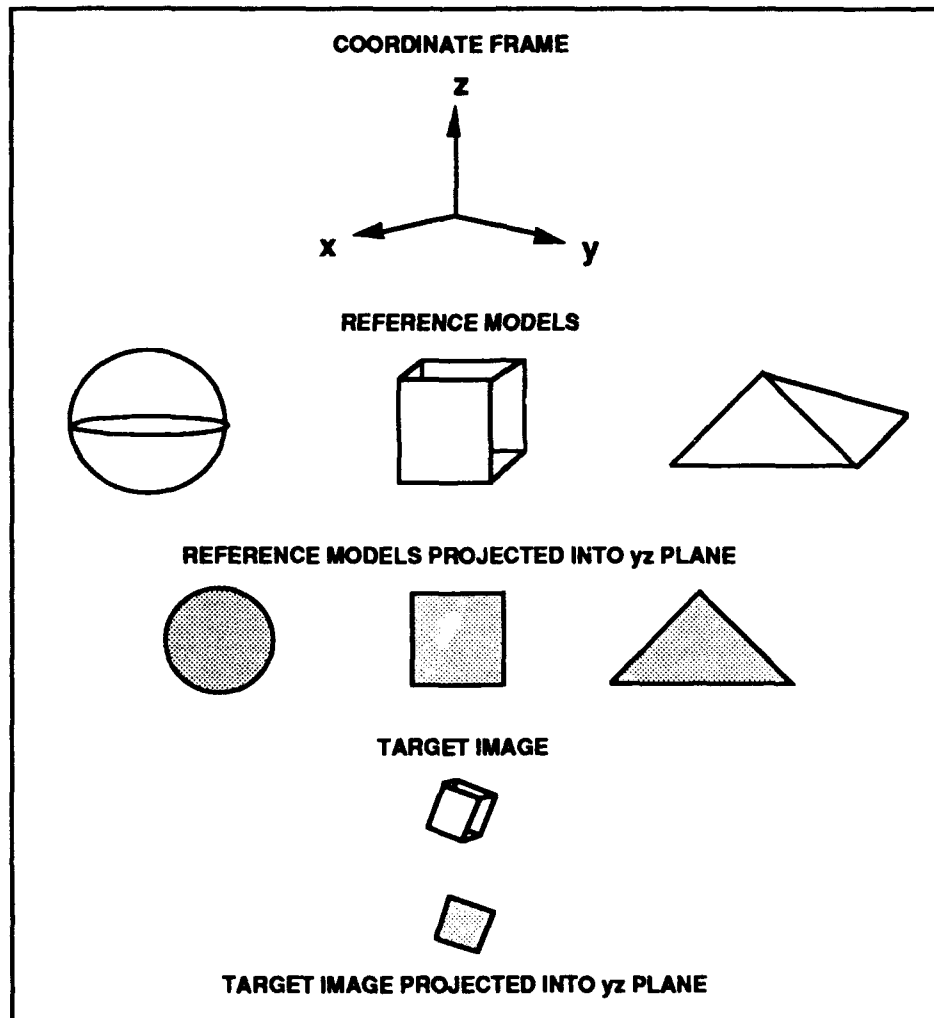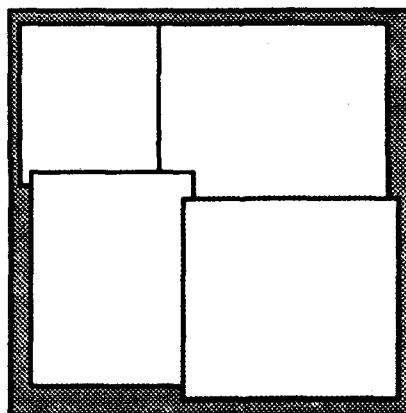
FIGURE 2. THREE-DIMENSIONAL ATR COMPARISON PROBLEM

ORIGINAL IMAGE WITH COVERING PIECES

FIGURE 3. COLLAGE PROCESS ASSOCIATED WITH ITERATED
FUNCTION SYSTEM IMAGE COMPRESSION PROCESS

VISION BACKGROUND

To solve the generic misalignment problem, one must be able to compute the local misalignments between a reference object and a transformed copy of itself (see Figure 4). For this discussion, the transformed copy is obtained from the reference object by the application of an affine transformation A(t). Simple processing cells with a sensitivity to image gradients across their field of view are appropriate for this task. This type of cell has appeared previously in the literature as part of boundary detection/completion systems (Reference 3).

These gradient detection cells possess the advantages of being similar to cells in the primate visual cortex and also easily implemented in terms of simple artificial neural network processors. By using an appropriate choice of the cell connection template, these cells may be adapted for analog very large scale integration (VLSI) implementation. The response of these cells would be recovered in a manner analagous to the *silicon retina* of Carver Mead (Reference 4).
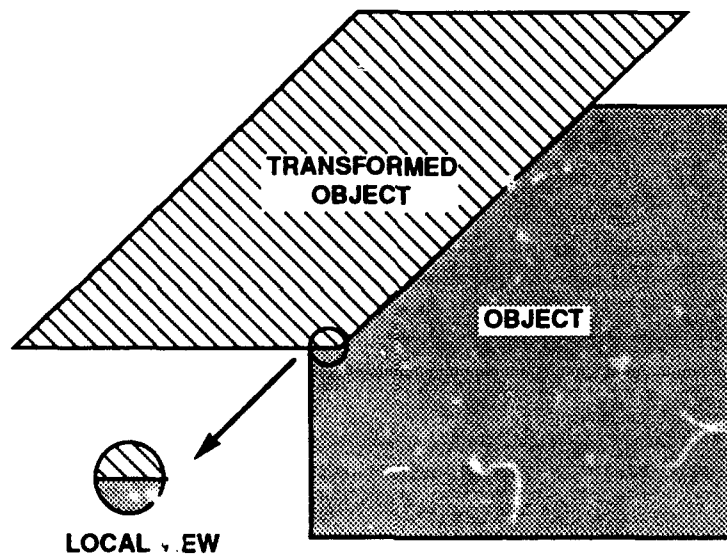
FIGURE 4. GENERIC MISALIGNMENT PROBLEM IN R^2

APPROACH

SIMPLE CELLS

Each side of a given simple cell may reside either in the object or transformed object pixel space. By varying the location of the *left* and *right* side of the simple cells between the object and transformed object, pixel images various types of configurations can be detected. A simple cell with both sides residing in the object at pixel location (i,j) is illustrated in Figure 5. This cell responds with an activation of 1 when

$$(L_0(i,j;0) - R_0(i,j;0)) > \alpha$$

where $L_0(i,j;0)$ is the integrated image intensity in the *left* side of this cell centered at (i,j) with orientation number 0 and is defined by

$$L_0(i,j;0) = \frac{\displaystyle\iint_{\text{left}} I_0(x,y)\, dx\, dy}{\displaystyle\iint_{\text{left}} dx\, dy}$$

R_0 is defined analogously, and α is a tolerance parameter. This cell is tuned to the type of local arrangement of the object in pixel space that is illustrated in Figure 5. The simple cells check for gradients along the four orientations 0, 45, 90, and 135 deg (Figure 6).
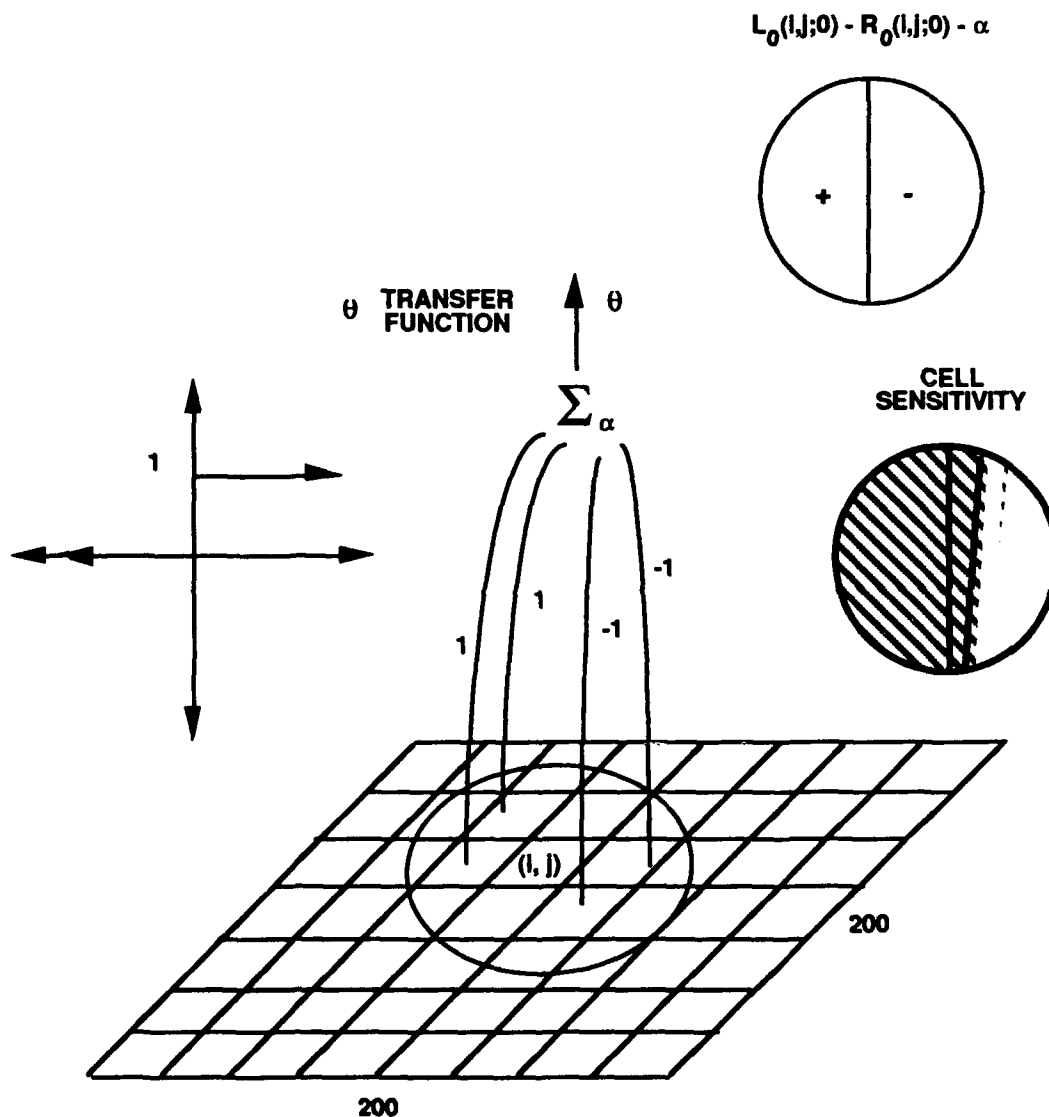
FIGURE 5. SIMPLE CELL PROCESSORS

ORIENTATION NO. 0
0 DEG

ORIENTATION NO. 1
45 DEG

ORIENTATION NO. 2
90 DEG

ORIENTATION NO. 3
135 DEG
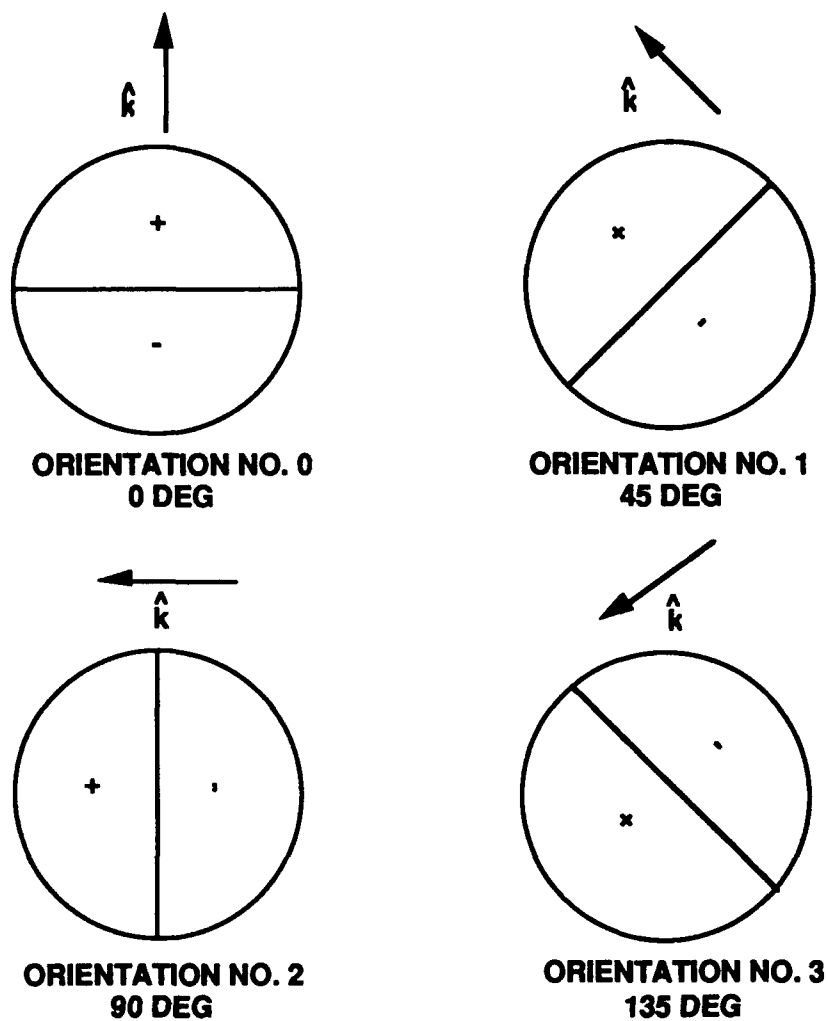
FIGURE 6. SIMPLE CELL ORIENTATIONS

By changing the sign between the *right* and *left* terms, one may vary the direction of cell sensitivity; or by allowing heterogeneous cells that respond to one sided gradients between the two pixel spaces in the same areas of the two images, misalignments between the object and the transformed object can be detected. The fundamental types of local arrangements require eight simple cell types as indicated in Figures 7 and 8. Four types are homogeneous in that both cell sides lie in the same pixel space, and four types are heterogeneous.

FOUR ORIENTATIONS: | — / \\

TYPE I – Θ $(L_0 - R_0 - \alpha)$

OBJECT

TYPE III – Θ $(R_0 - L_0 - \alpha)$

OBJECT

TYPE II – Θ $(L_{TO} - R_{TO} - \alpha)$

TRANSFORMED OBJECT

TYPE IV – Θ $(R_{TO} - L_{TO} - \alpha)$

TRANSFORMED OBJECT
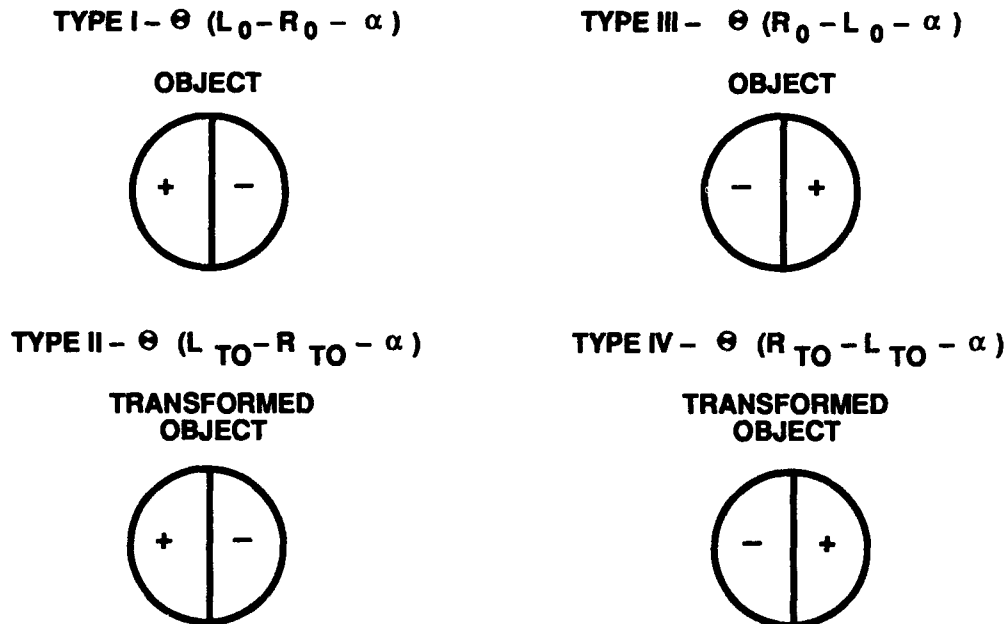
FIGURE 7. SIMPLE PROCESSORS (HOMOGENEOUS)

TYPE V – Θ (L$_0$ – L$_{TO}$ – β)
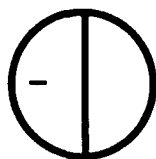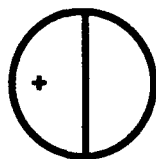
OBJECT

TRANSFORMED
OBJECT

TYPE VII – Θ (R$_0$ – R$_{TO}$ – β)

OBJECT

TRANSFORMED
OBJECT

TYPE VI – Θ (L$_{TO}$ – L$_0$ – β)
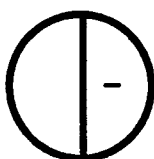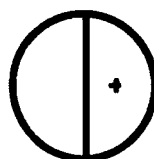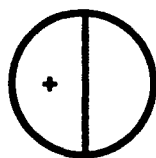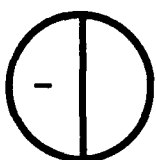
OBJECT

TRANSFORMED
OBJECT

TYPE VIII – Θ (R$_{TO}$ – R$_0$ – β
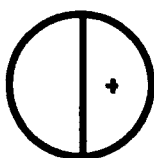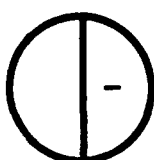
OBJECT

TRANSFORMED
OBJECT

FIGURE 8. SIMPLE PROCESSORS (HETEROGENOUS)

COMPLEX CELLS

The simple cells may be combined together to create complex cells. Based on the type of mismatch between the object and the transformed object, these complex cells indicate the local correction needed in the transformed objects position. These complex cells perform a logical "and" operation on the outputs of the simple cells that represent the salient features of the configuration of the object/transformed object pair. If the complex cell responds with an activation of 1, this indicates the correction needed on the transformed object to improve its alignment with the reference object (Figure 9).

OBJECT

TRANSFORMED OBJECT

NEEDED CORRECTION ON TRANSFORMED OBJECT

M

$$M = \begin{cases} 1 & \longrightarrow \\ 0 & \text{NO MISALIGNMENT OF THIS TYPE} \end{cases}$$

π

1   1   1   1

$\Theta (L_0 - L_{TO} - \beta)$ $\Theta (R_0 - R_{TO} - \beta)$ $\Theta (L_{TO} - R_{TO} - \alpha)$ $\Theta (L_0 - R_0 - \alpha)$

NOTE: Simple Cell Orientation Vector is K
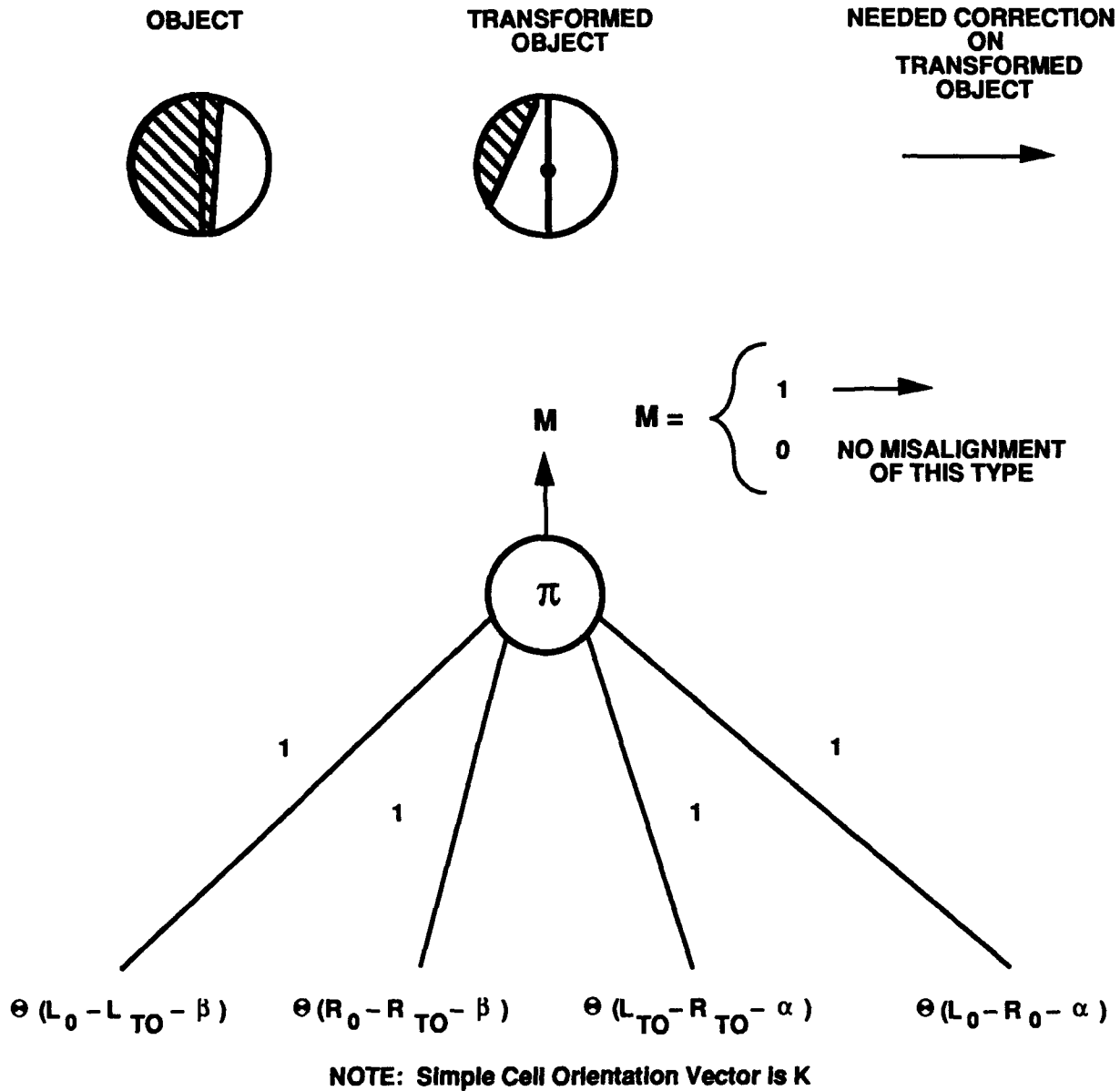
FIGURE 9. COMPLEX CELL ARCHITECTURE

For a given orientation, there are four object/transformed object cases and four corresponding complex cells. These configurations have been designed to respond to misalignments along the borders of the sets. Two of these cells indicate a correction on the transformed object in the k direction, and two of these cells indicate a correction in the -k direction. As indicated in Figure 10, these cells may be wired into a sigma unit in such a manner that the net output indicates not only whether a local correction along this orientation is needed, but also the direction of the correction. For each point in pixel space, a value of +1, 0, -1 is assigned for each of the four orientations. These are called the local gradients and are denoted at time t, orientation n, and pixel location (i,j) as $\gamma_{ijn}(t)$.

OBJECT

TRANSFORMED
OBJECT

COMPLEX CELL
NO. 1

COMPLEX CELL
NO. 2

COMPLEX CELL
NO. 3

COMPLEX CELL
NO. 4

$$M = \begin{cases} +1 & \longrightarrow \\ -1 & \longleftarrow \\ 0 & \text{NO CHANGE} \end{cases}$$

M

Θ

Σ

$\Theta \equiv$

+1

−1

+1

−1

−1

+1

CC1

CC2

CC3

CC4
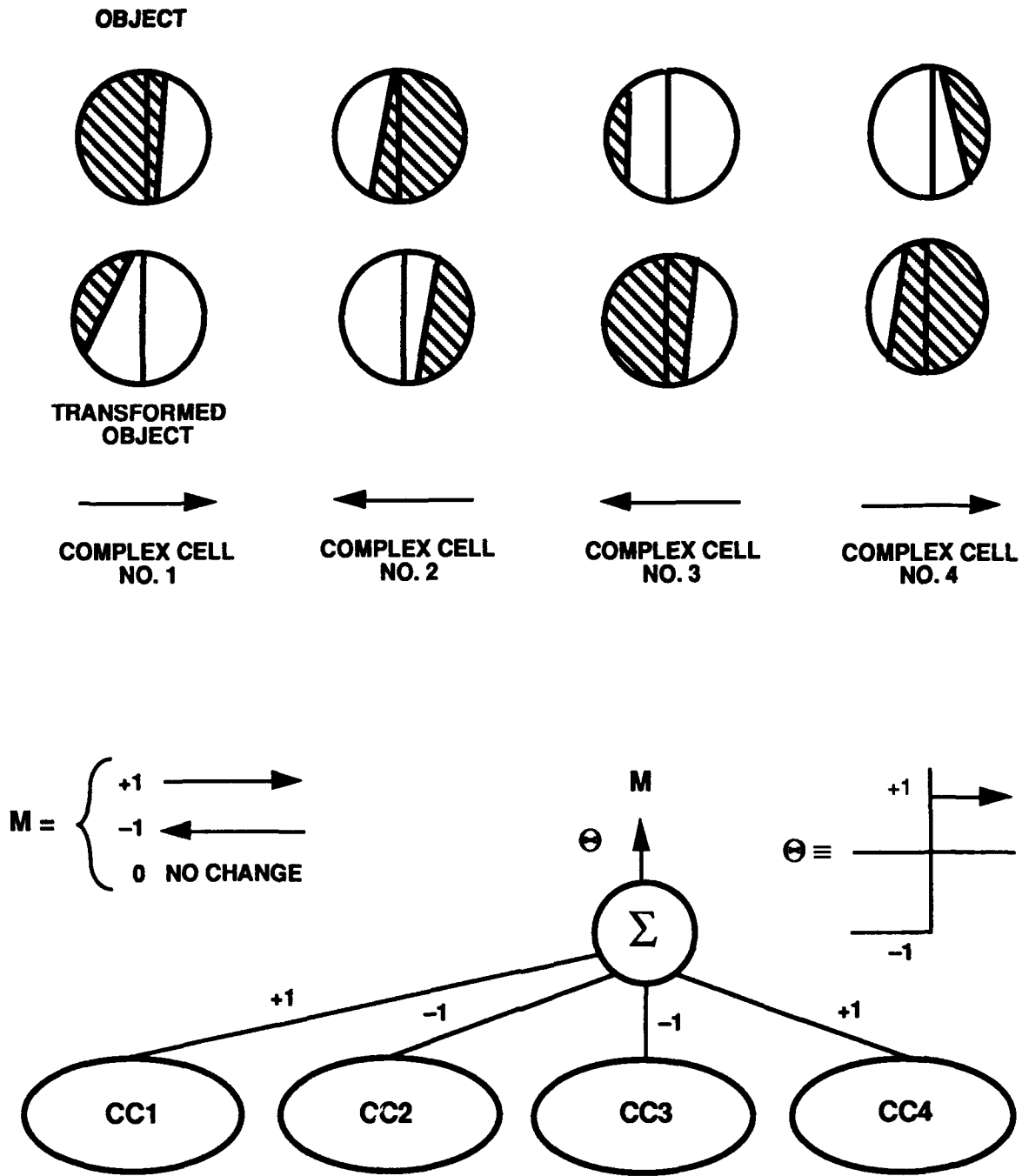
FIGURE 10. TYPES OF MISALIGNMENTS

Figure 11 portrays the four local gradient planes for a representative sample case at time t=0. A symbol has been plotted at those points in pixel space where the local gradients are nonzero. As portrayed in the figure, the local gradients represent the the response of the complex cells to boundary misalignments between the object and its transformed copy. As expected, there is no response on the overlapping interiors of the two objects and on those mismatched boundary sections that exceed the diameter of the simple cells.
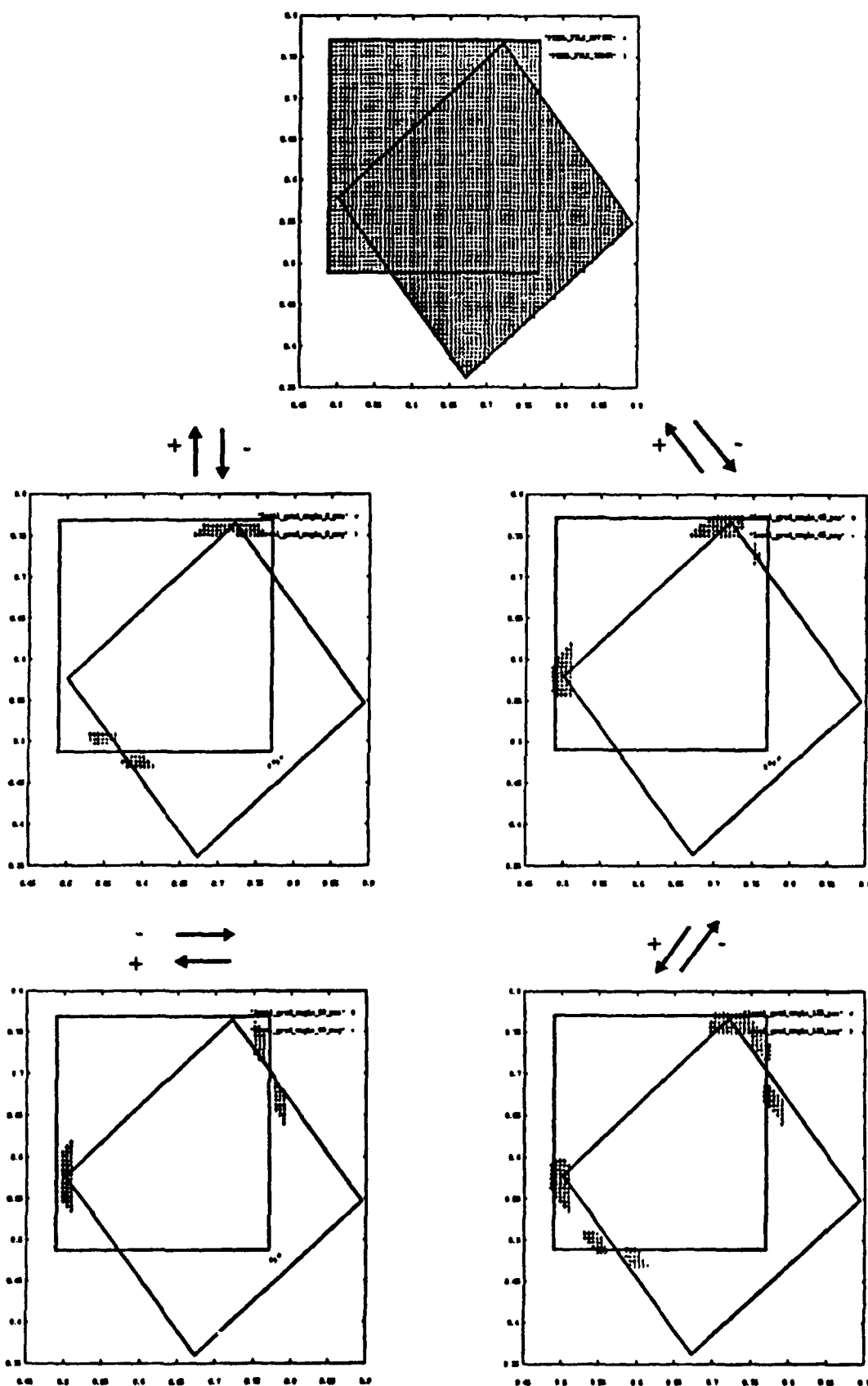
FIGURE 11. LOCAL GRADIENTS

12

GAUSSIAN AVERAGING

The six parameters that determine an affine mapping are uniquely determined by the action of the mapping on three points. Therefore, three points residing in the transformed object image inust be chosen in order to find the new affine transformation that will improve the overlap between the transformed and reference object. These three points are chosen close to the boundary of the transformed object to make best use of the local gradients. The first point is chosen along the ray connecting the center of mass of the transformed object to the point of the set that is at a maximal distance from the center of mass. The other two points are picked equally distributed in angular space relative to this ray (see Figure 12).
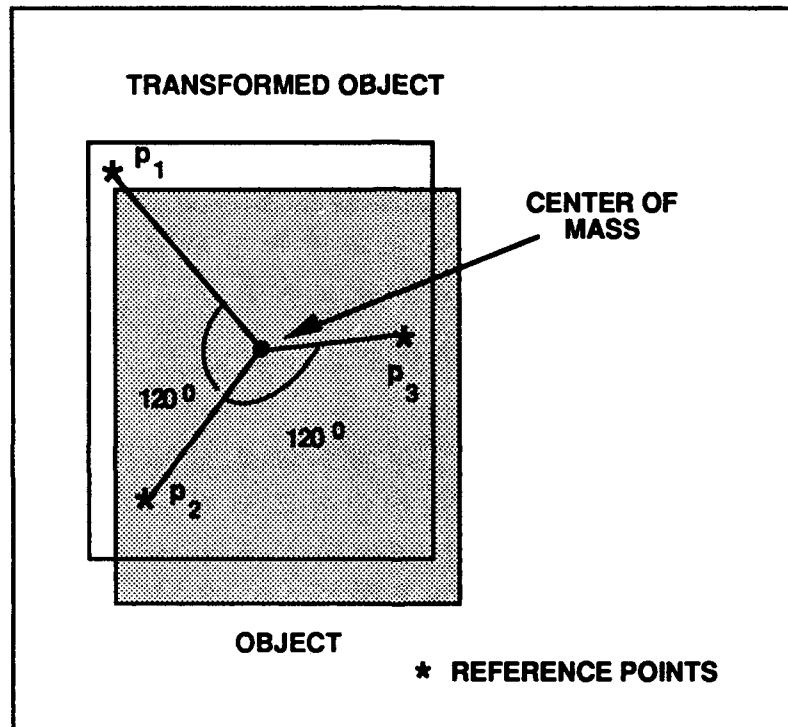
FIGURE 12. REFERENCE POINTS USED FOR GAUSSIAN
GRADIENT COMPUTATION

Let the three points be designated p1=(i1, j1), p2= (i2, j2), and p3=(i3, j3). For each of the three points (iλ,jλ) λ=1..3, the goal is to compute a desired correction vector $\Delta r_\lambda(t)$ based on the influence of the local gradients. The local gradients contribute a Gaussian weighted term to the $\Delta r_\lambda(t)$ value at each point as part of a global averaging process. Let d be the squared Euclidean distance between pixels (iλ, jλ) and (i,j), and let d0 be 0.25 times the squared length of the diagonal of the rectangle that contains the transformed object. The Gaussian gradient at point (iλ, jλ), time t, and orientation n is given by

$$\Gamma_n(i\lambda, j\lambda; t) = \sum_i \sum_j \gamma_{ijn}(t) \exp[-d/d0]$$

Once these are computed, we may compute the $\Delta r_\lambda(t)$ using the vector components of the Gaussian gradients

$$\text{x component of } \Delta r_1(t) = \sum_{n=0}^{3} \hat{k}_n * \hat{i} * \Gamma_n(il, jl; t)$$

$$\text{y component of } \Delta r_1(t) = \sum_{n=0}^{3} \hat{k}_n * \hat{j} * \Gamma_n(il, jl; t)$$

Similarly for $\Delta r_2(t)$ and $\Delta r_3(t)$. Once obtained, each of the $\Delta r_\lambda(t)$ is normalized separately.

COMPUTATION OF NEW TRANSFORM

The new points $p1'$, $p2'$, $p3'$ are computed using

$$p\lambda' = p\lambda + \Delta r_\lambda(t) * \text{rstep}$$

where rstep is the current step size being used in the gradient descent process.

The transform $T(t)$ which takes the $p\lambda$ to the $p\lambda'$ is recovered by solving

$$\begin{vmatrix} i_1 & j_1 & 1 \\ i_2 & j_2 & 1 \\ i_3 & j_3 & 1 \end{vmatrix} \begin{bmatrix} a & c \\ b & d \\ e & f \end{bmatrix} = \begin{vmatrix} i_1'' & j_1'' \\ i_2'' & j_2'' \\ i_3'' & j_3'' \end{vmatrix}$$

Finally the new A transformation at time t+1 is computed

$$A(t+1) = T(t) \, o \, A(t)$$

where o indicates right to left functional composition. This new transformation improves the alignment of the boundaries; and therefore, the interiors of the transformed and reference object.

NETWORK ARCHITECTURE

The complex cells for the four orientations may be combined together along with their simple cell building blocks to produce a local gradient network architecture as portrayed in Figure 13. The network inputs the pixel responses into each simple cell and combines the simple

cells together to form each complex cell. Finally, the network fuses the four complex cell activities for each orientation to produce the four local gradients M_1, M_2, M_3, and M_4 at each point.

The Gaussian gradient calculation can also be performed in a connectionist manner. Given the transformation A(t), the Gaussian gradient terms can be viewed as connections between two fully interconnected pixel planes. With this convention in mind, it is possible to compute the total number of connections in the network. Assuming a simple cell radius of 6, there are approximately 100 interconnections per simple cell. Using this value, we can compute the number of interconnections per pixel in the local gradient portion of the network as

$$(100 \text{ intcon/sc})(4\text{sc/cc})(4\text{cc/orient.})(4\text{orient})(\text{\# of pixels}) = 6400$$

For a 200 by 200 pixel space, this produces around 256 million interconnection for the local gradient portion of the network. The Gaussian gradient values require another $(200)(200)(200)(200) = 1.6 \times 10^9$ interconnects for storage; however, only $(3)(200)(200)$ of these are active at any given time. This produces a total of 1.8 billion interconnects, of which 256 million are active at any given time.

LOCAL GRADIENTS AT PIXEL (I,J) FOR EACH ORIENTATION
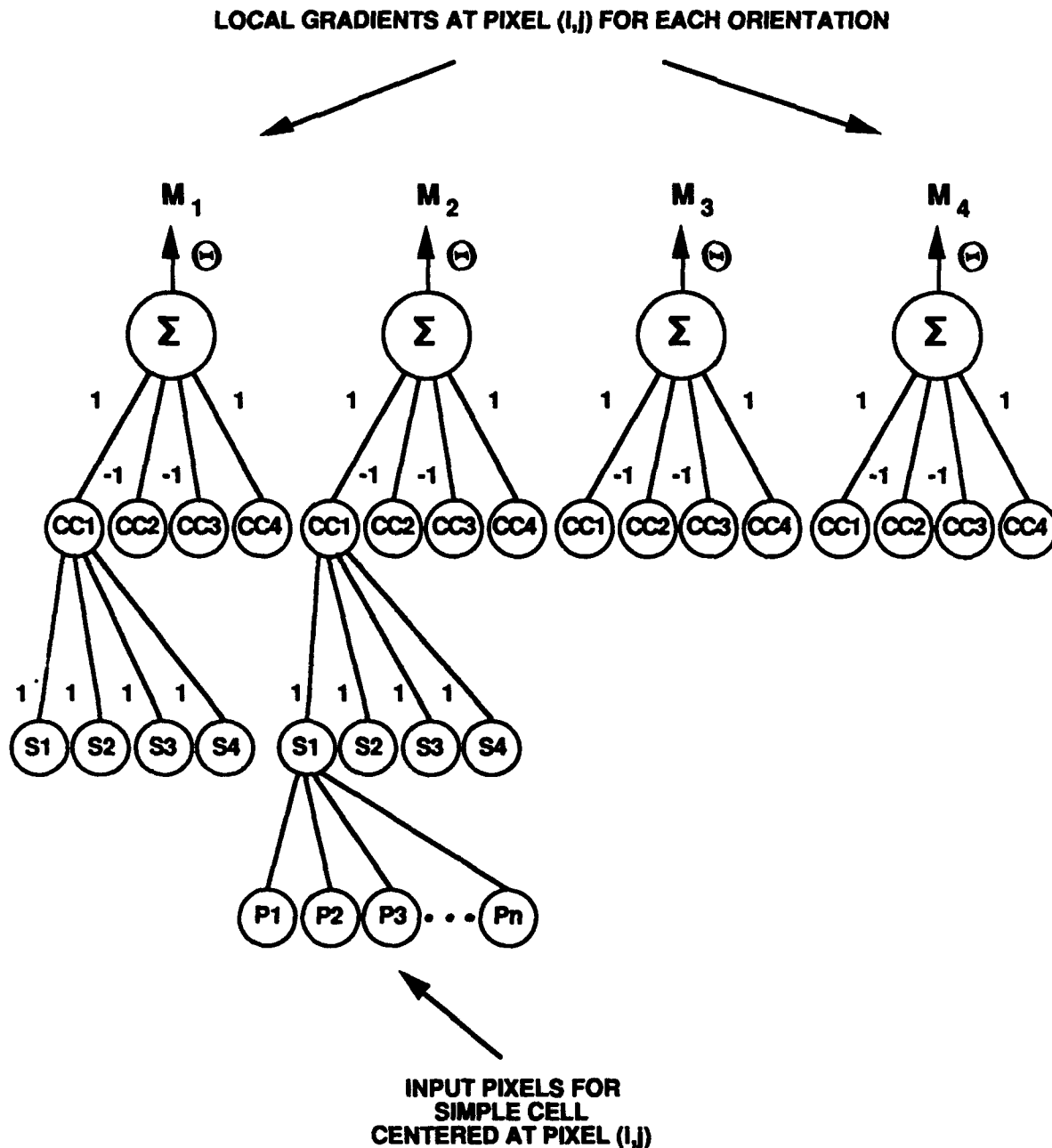
FIGURE 13. NETWORK STRUCTURE FOR LOCAL GRADIENT COMPUTATION AT PIXEL (i,j)

CONVERGENCE PROPERTIES FOR ONE–DIMENSIONAL CASE

The convergence properties of the VGD network will be illustrated with the analysis of two simple one-dimensional cases. The one-dimensional general affine transformation is of the form $f(x) = a(x) + b$, where a and b $\in$ R. In this case, the affine transformation is uniquely determined by its action on two points.

CASE 1–TRANSFORMATION IS SIMPLE TRANSLATION

Let $r=1$, $0<\alpha=\beta=\gamma<<1$. Consider the simple case where the object set is I=[0,1] , E=[a,1+a] is the transformed object, and $\gamma <a<1- \gamma$. That is, E is obtained from I by the action of a pure translation (to the right, without loss of generality) of I by less than the cell radius r, and I $\cap$ E $\neq \emptyset$. The region ACR[1] of interest (the only points that can possibly be active for a cell centered at x$\in$ E) is A=[a-r,1+a+r].

There are five subregions to consider:

$$A = \cup A_i \ (i=1,..,5) = [a-r,0] \cup [0,a] \cup [a,1] \cup [1,1+a] \cup [1+a,1+a+r].$$

We wish to show that the total contribution of restorative cells

$$\rho_1 = \{x\in A : \text{Case 2 or Case 3 holds}\}$$

is greater than any possible contribution from improper motion

$$\rho_2 = \{x\in A : \text{Case 1 or Case 4 holds}\}$$

thereby *producing restorative dynamics*. Case here refers to the object-transformed object cases illustrated in Figure 14. To simplify notation, the Lebesgue measure of the left side of the one dimensional simple cell residing in the set I will be denoted L(I) instead of $\mu(L_I)$. For x$\in$ A$_1$ = [a-r,0] we have L(I) = L(E) = 0, ruling out all four cases. Thus, $\rho_1 \cap A_1 = \rho_2 \cap A_1 = \emptyset$. Similarly, for A$_5$ = [1+a,1+a+r], $\rho_1 \cap A_5 = \rho_2 \cap A_5 = \emptyset$. For x$\in$ A$_3$=[a,1] we have

L(I) = x
R(I) = 1-x
L(E) = x-a
R(E) = 1-(x-a)

In particular, R(E) > R(I), ruling out Cases 1 and 2, and L(I) > L(E), ruling out Cases 3 and 4. Thus, $\rho_1 \cap A_3 = \rho_2 \cap A_3 = \emptyset$. It remains only to consider A$_2$ = [0,a] and A$_4$ = [1,1+a]. For x$\in$ A$_2$, we have

L(I) = x
R(I) = 1-x
L(E) = 0
R(E) = 1-(a-x)

L(E) = 0 rules out Cases 1, 3, and 4. The conditions for Case 2 are

(i) $x > \gamma$

(ii) $1-x > 1 - (a-x) + \gamma$

(iii) $1 - (a-x) > \gamma$

(iv) $1-x > x + \gamma$

(i) and (iii) are satisfied for $x > \gamma$. (ii) and (iv) imply $(x < 1/2 - \gamma/2) \cap (x < a/2 - \gamma/2)$. That is, Case 2 is satisfied for

$$\Delta_1 = \{x : \gamma < x < \min(1/2 - \gamma/2 , a/2 - \gamma/2)\}$$

For $x \in A_4$, we have (writing $x = 1+y$, $y \in [0,a]$)

$L(I) = 1-y$
$R(I) = 0$
$L(E) = 1 - (a-y)$
$R(E) = a-y$

$R(I) = 0$ rules out Cases 1, 2, and 4. The conditions for Case 3 imply $(a-y < 1/2 - \gamma/2) \cap (y > a/2 + \gamma/2)$, or Case 3 is satisfied for

$$\Delta_2 = \{y : \max(a - 1/2 + \gamma/2 , a/2 + \gamma/2) < y < a - \gamma \}$$
$$= \{x : \max(a + 1/2 + \gamma/2 , a/2 + 1 + \gamma/2) < x < a +1 - \gamma \}$$

In summation, then, we have $\rho_1 \cap [A_2 \cup A_4] > \rho_2 \cap [A_2 \cup A_4] = \emptyset$. Thus we have net restorative action, as desired. For convergence (that is, $E \to I \pm g$), it suffices then to consider the step size $s_n \to 0$, with $\Sigma n s_n = \infty$. For recovery of transformation, we may consider having chosen one point from each of Δ_1, Δ_2. That is, $x_1 \in \Delta_1$, $x_2 \in \Delta_2$ will both yield restorative action. Hence, overall action will be translation to the left as desired. An obvious choice, a priori, for x_1, x_2 is $x_1 = a$, $x_2 = 1+a$. This choice will assure the maximum Gaussian contribution to the restorative force from those points in Δ_1 and Δ_2.

LOCAL VIEWPOINT   **GLOBAL VIEWPOINT**

CASE 1

$[\mu(L_I) - \mu(L_E) > \beta] \wedge [\mu(R_I) - \mu(R_E) > \beta] \wedge [\mu(L_E) - \mu(R_E) > \alpha] \wedge [\mu(L_I) - \mu(R_I) > \alpha]$

CASE 2

$[\mu(L_I) - \mu(L_E) > \beta] \wedge [\mu(R_I) - \mu(R_E) > \beta] \wedge [\mu(R_E) - \mu(L_E) > \alpha] \wedge [\mu(R_I) - \mu(L_I) > \alpha]$

CASE 3

$[\mu(L_E) - \mu(L_I) > \beta] \wedge [\mu(R_E) - \mu(R_I) > \beta] \wedge [\mu(L_I) - \mu(R_I) > \alpha] \wedge [\mu(L_E) - \mu(R_E) > \alpha]$

CASE 4

$[\mu(L_E) - \mu(L_I) > \beta] \wedge [\mu(R_E) - \mu(R_I) > \beta] \wedge [\mu(R_I) - \mu(L_I) > \alpha] \wedge [\mu(R_E) - \mu(L_E) > \alpha]$

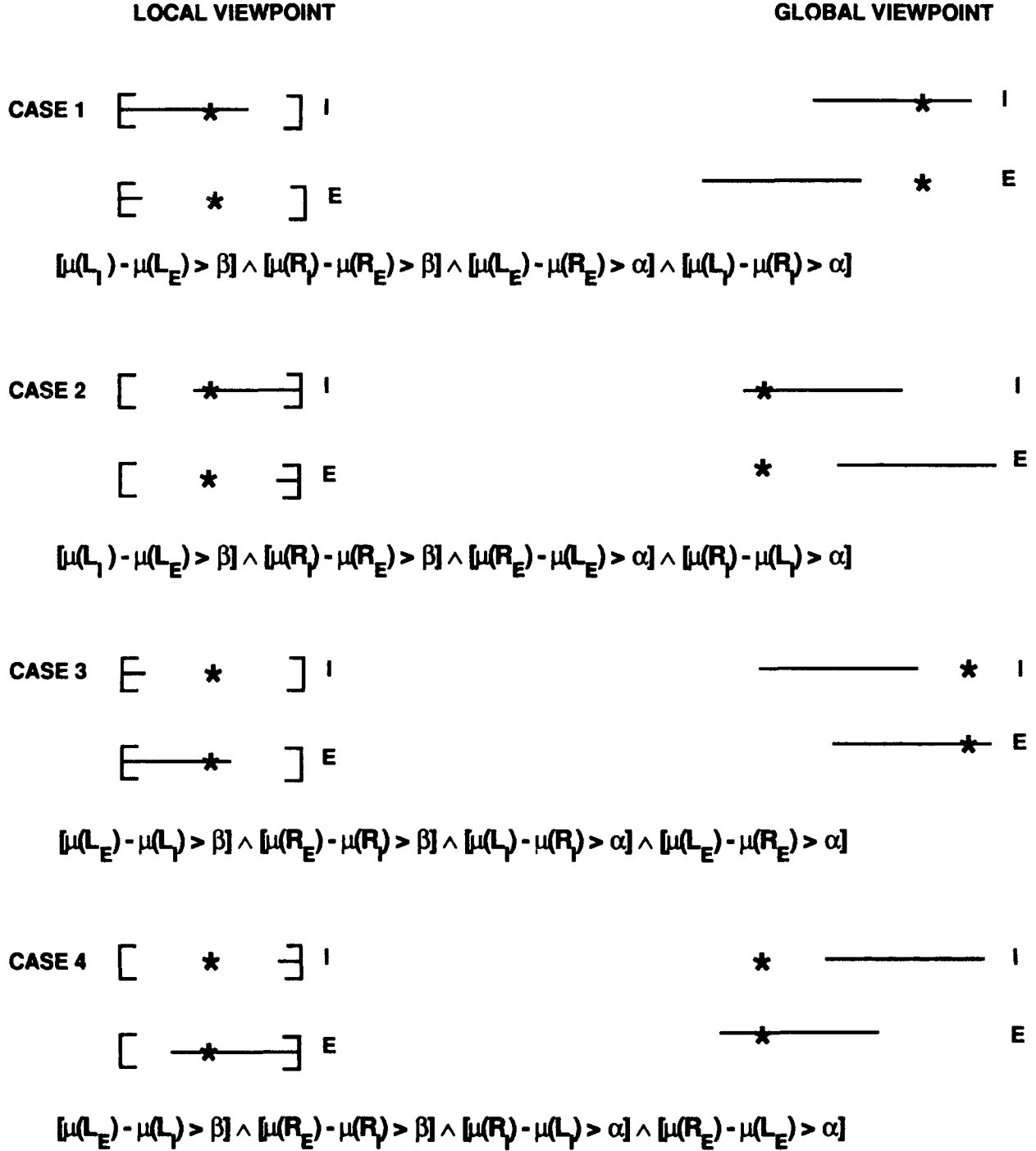FIGURE 14. ONE-DIMENSIONAL OBJECT TRANSFORMED OBJECT CASES

CASE 2–TRANSFORMATION IS SIMPLE SCALING

Again, let $r=1$, $0<\alpha=\beta=\gamma<<1$. Consider $I=[0,1]$, $E=[0,a]$, $\gamma <a<1-\gamma$. That is, E is a simple scaling of I. Then, as before, we write the region of interest A as A = $[-r,a+r]$ = $\cup A_i$ (i=1,...,4),

19

where $A_1=[-r,0]$, $A_2=[0,a]$, $A_3=[a,1]$, and $A_4=[1,1+r]$. Similar to above, we obtain $\rho_1 \cap A_1 = \rho_1 \cap A_4 = \rho_2 \cap A_1 = \rho_2 \cap A_4 = \emptyset$. Thus it remains to consider A_2 and A_4. For $x \in A_2$, $L(I) = L(E) = x$, ruling out all four cases, and $\rho_1 \cap A_2 = \rho_2 \cap A_2 = \emptyset$. For $x \in A_3$, we have

$$L(I) = x$$
$$R(I) = 1-x$$
$$L(E) = a$$
$$R(E) = 0$$

Case 1 is satisfied for

$$\Delta_3 = \{x : \max(1/2 + \gamma/2 , a + \gamma) < x < 1 - \gamma\}$$

Thus, $\rho_2 \cap A_3 > \rho_1 \cap A_3 = \emptyset$, and we have net restorative effect. Convergence then requires only the conditions on the step size s_n noted above. For transformation recovery, we see that $x_1=0$, $x_2=a$ again yield the required dynamics.

For a case in which we have both scaling and translation, the translation effects align the objects (to within γ of perfect alignment), then the scaling takes place, and convergence is maintained.

RESULTS

The capability of the VGD network to recover an unknown affine transformation can best be illustrated with a simple example. The results that are presented here were produced by a serial implementation of the VGD network running on a Silicon Graphics 4D/220. One would expect the results produced by an analog implementation of the network architecture to be similar.

In the standard ATR process, a transformation is sought that optimally aligns a reference model with a copy of itself. In this example, the equivalent problem of aligning the transformed copy with the reference model is solved. Given a copy of a square that has been subjected to an unknown affine transformation, we wish to solve for the inverse transform.

Figure 15 portrays the 50 steps needed by the VGD process to recover the transformation. The upper left corner of the figure represents the initial configuration of the square and its transformed copy. The salient simulation parameters for the run are summarized in Table 1. The first 20 iterations act to align the centers of the two images. The last 30 iterations scale the transformed copy and rotate it into position over the reference image. After 50 iterations, the 2 images are aligned to a tolerance of 1 pixel. The number of steps required for convergence is small compared to the 500 or so needed steps required by a random search technique such as generalized simulated annealing (Reference 5).

INITIALIZATION

10 ITERATIONS

20 ITERATIONS

30 ITERATIONS

40 ITERATIONS

50 ITERATIONS
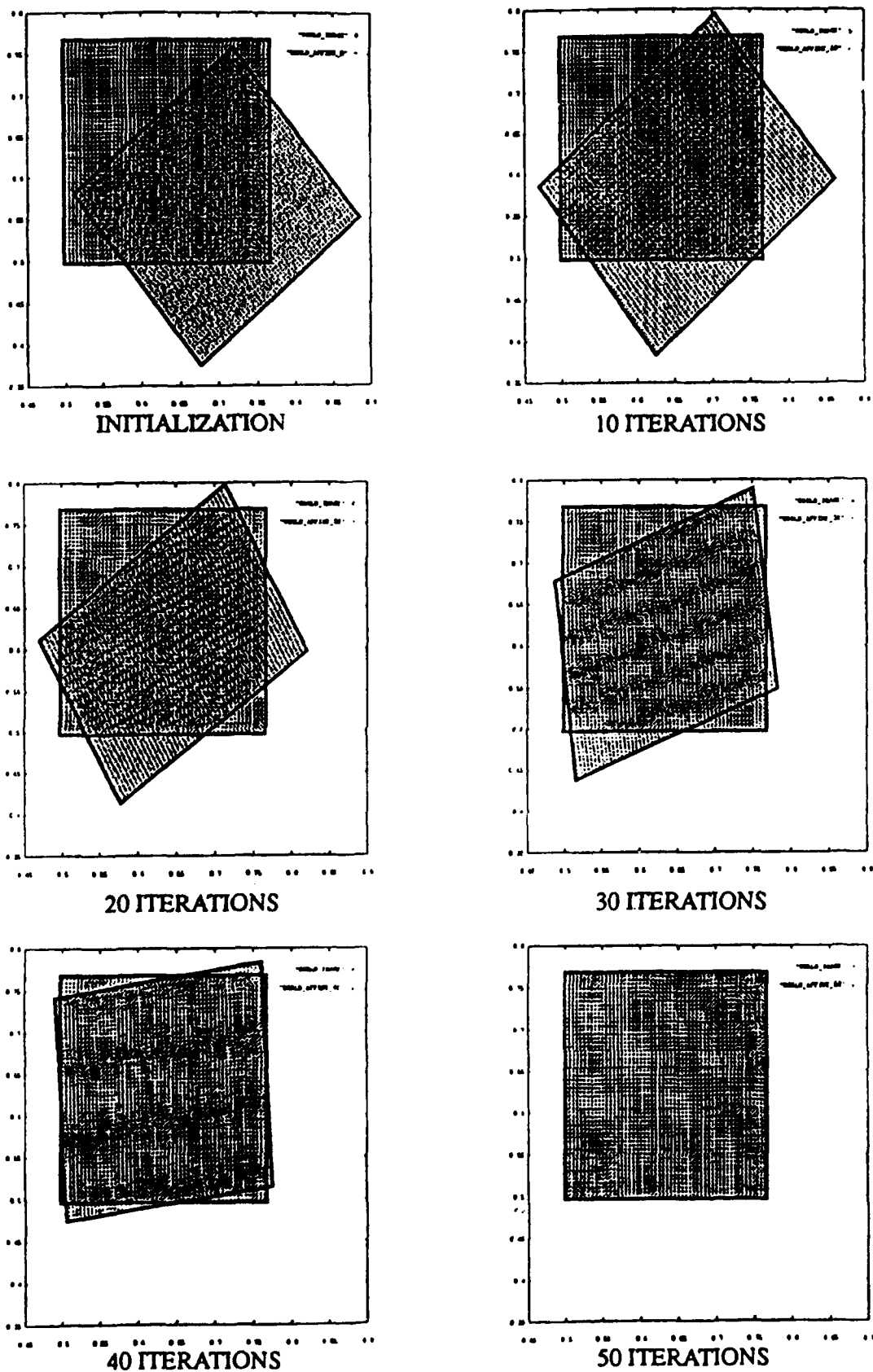
FIGURE 15. VGD RESULTS FOR SQUARE

TABLE 1. SIMULATION PARAMETERS

CHARACTERISTIC	MEASUREMENT
Image Limits	-.05 to 1.0
Image Size	200 x 200 pixels
Simple Cell Radius	5 pixels
Cell Threshhold	3 pixels
Step Size	.005 (1 pixel)

CONCLUSION

The VGD network provides a new technique to rapidly recover an unknown transformation relating two objects. Depending on the type of transformation, the VGD network has direct applications to two- and three-dimensional ATR and to image compression using IFS. Although not impervious to local minima, the rapid convergence of this guided technique offers clear advantage over random search techniques such as simulated annealing and genetic algorithms.

The authors are continuing work on the use of the VGD network. Current work includes a parallel version of the VGD network that covers an image with multiple copies of itself, a model-based 6 DoF ATR system employing both generalized simulate annealing and VGD and an analog implementation of the VGD network. All of these efforts provide further evidence of the utility and power of the generic VGD architecture.

REFERENCES

1.   Goldstein, H. (1981), *Classical Mechanics*, Reading: Addison-Wesley.

2.   Barnsley, M. (1989). *Fractals Everywhere*. San Diego: Academic Press.

3.   Carpenter, G.; Grossberg, S.; and Mehanian, C., "Invariant Recognition of Cluttered Scenes by a Self- Organizing ART Architecture: CORT-X Boundary Segmentation," *Neural Networks*, No. 2, 1989.

4.   Maher, M.; Deweerth, M; Mahowald. M.; and Mead, C., "Implementing Neural Architectures Using analog VLSI Circuits," *IEEE Transactions on Circuits and Systems*, 1989.

5.   Solka, J.; Vermillion, D.; and Rogers, G., "Simulated Annealing and Iterated Function Systems: A Collaboration for Data Reduction," *Computer Vision, Graphics, and Image Processing*, in review.

DISTRIBUTION

	COPIES
ATTN CODE 213 SIEGEL	1
OFFICE OF NAVAL TECHNOLOGY	
800 N QUINCY ST	
ARLINGTON VA 22217-5000	
ATTN THOMAS MCKENNA	1
OFFICE OF CHIEF OF NAVAL RESEARCH	
COGNITIVE AND NEURAL SCIENCE DIVISION	
800 N QUINCY ST	
ARLINGTON VA 22217-5000	
ATTN CLIFFORD LAU	1
OFFICE OF CHIEF OF NAVAL RESEARCH	
ELECTRONICS DIVISION	
800 N QUINCY ST	
ARLINGTON VA 22217-5000	
ATTN DSO YOUN	1
DEFENSE ADVANCED RESEARCH	
PROJECTS AGENCY	
1400 WILSON BLVD	
ROSLYN VA 22201	
ATTN PAUL WERBOS	1
NATIONAL SCIENCE FOUNDATION	
ROOM 1151	
WASHINGTON DC 20550	
DEFENSE TECHNICAL INFORMATION	
CENTER	2
CAMERON STATION	
ALEXANDRIA VA 22304-6145	
ATTN AL GORDON	1
COMMANDER	
NAVAL OCEAN SYSTEMS CENTER	
SAN DIEGO CA 92152-5000	
ATTN ROY STREIT	1
COMMANDER	
NAVAL UNDERWATER SYSTEMS CENTER	
NEW LONDON CT 06320	

	COPIES
ATTN JIM MAAR R51	1
NATIONAL SECURITY AGENCY	
FT MEADE MD 20755-6000	
ATTN SEA 06	1
PMS 400	1
COMMANDER	
NAVAL SEA SYSTEMS COMMAND	
WASHINGTON DC 20362-5101	
ATTN OP 03	1
OP 033	1
OP 035	1
OP 037	1
CHIEF OF NAVAL OPERATIONS	
DEPARTMENT OF THE NAVY	
WASHINGTON DC 20360	

INTERNAL DISTRIBUTION

Code	Name	COPIES
C		1
D		1
D4		1
E211	FINK	1
E231		3
E232		2
E32	GIDEP	1
E42	HYNSON	1
F33	TURNER	1
F44	POWLAK	1
G11	LUCAS	1
G23	MALYEVAC	1
G42	FARSAIE	1
G42	DOMOULIN	1
G43	CAVARIS	1
H23	DRAPER	1
K		1
K10		5
K12		5
K12	ROGERS	10
K12	POSTON	2
K12	PRIEBE	10
K12	SOLKA	10
K13		1
K14		1
K43	VERMILLION	10
K44	GLASS	1
K44	REID	1

DISTRIBUTION (CONTINUED)

K53	COOPER	1
N35	FENNEMORE	1
N35	KUCHINSKI	1
N35	TAFT	1
N35	TATE	1
N35	FENNEMORE	1
N415	MCCLINTOCK	1
N415	HARMAN	1
R04	RIEDL	1
R301	JONES	1
R44	SZU	1
R44	TELFER	1
R44	BARNETT	1
U042	BEILMANIS	1
U23	BARAN	1

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE November 1991	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE
Connectionist Approach to Transformation Recovery Using Visual Gradient Descent

5. FUNDING NUMBERS

6. AUTHOR(S)
George W. Rogers, Jeffrey L. Solka, Donald R. Vermillion, Carey E. Priebe

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Surface Warfare Center (K12)
Dahlgren, Virginia 22448-5000

8. PERFORMING ORGANIZATION REPORT NUMBER
NAVSWC TR 91-609

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Given an object and a copy of itself produced by an unknown two-dimensional affine transformation, a new neural network architecture has been developed that recovers this transformation by minimizing the symmetric difference between the object and the copy. This architecture performs a gradient descent in symmetric difference error space and is designated as visual gradient descent (VGD). The VGD network has applications to both two- and three-dimensional model based automatic target recognition (ATR) and image compression using iterated function systems.

14. SUBJECT TERMS
Neural Network, Gradient Descent, Symmetric Difference, Automatic Target Recognition (ATR), Image Compression, Iterated Function System, Local Gradients, Optimization.

15. NUMBER OF PAGES
20

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and its title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet *optical scanning requirements.*

Block 1. Agency Use Only *(Leave blank).*

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	-	Contract	PR	- Project
G	-	Grant	TA	- Task
PE	-	Program Element	WU	- Work Unit Accession No.

BLOCK 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If Known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD	-	See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE	-	See authorities.
NASA	-	See Handbook NHB 2200.2
NTIS	-	Leave blank

Block 12b. Distribution Code.

DOD	-	Leave blank.
DOE	-	Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA	-	Leave blank.
NTIS	-	Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*

Block 17.-19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of this page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.